Recovering boundary conditions in inverse Sturm-Liouville problems

Norbert Röhrl

Abstract. We introduce a variational algorithm, which solves the classical inverse Sturm-Liouville problem when two spectra are given. In contrast to other approaches, it recovers the potential as well as the boundary conditions without a priori knowledge of the mean of the potential. Numerical examples show that the algorithm works quite reliable, even in the presence of noise. A proof of the absence of strict local minimizers of the functional supports the observation, that a good initial guess is not essential.

1. Introduction

The inverse Sturm-Liouville problem was first systematically studied by Borg in 1946 [**1**]. He already proved that all information needed to reconstruct the potential is in two sequences of eigenvalues, and applied this to the question if one could hear the mass density of a guitar string.

With modern computers, the question for efficient algorithms to actually compute the potential gained importance [**7**]. This work was inspired by two different approaches to this problem.

One was developed by Rundell and Sacks [**10**] and is based on the Gelfand-Levitan-Marchenko kernels. It is elegant and efficient, but also invariably needs the mean $\int_0^1 Q \, \mathrm{d}x$ of the potential and the boundary conditions as additional inputs besides the two spectra. The second method is variational and was created by Brown, Samko, Knowles, and Marletta [**2**]. It does not need the mean as a separate input, but it is unknown if it also can be used to recover the boundary conditions.

We also want to mention a related recovery method by Lowe, Pilant, and Rundell [**6**], which uses a finite basis ansatz, and solves the inverse problem by Newton's method without requiring the mean as input.

In this paper we extend the variational method we introduced in [**9**] to recover potential and boundary conditions in the case when only two finite sequences of eigenvalues are given. Having less reliable information it is not as robust under noisy input, but we still get reasonable results.

In the following section we will define the functional and exhibit some essential properties. The numerical examples are discussed in the third section and section four finally contains the proof of the absence of strict local minimizers.

2. Definition and Properties of the Functional

We consider the Sturm-Liouville equation

$$-u'' + q(x)u = \lambda u \tag{SL}$$

on $[0,1]$ with $q(x) \in L^2([0,1],\mathbb{R})$ real, and separated boundary conditions

$$h_0 u(0) + u'(0) = 0, \quad h_1 u(1) + u'(1) = 0\,. \tag{$h_0 h_1$}$$

2000 *Mathematics Subject Classification.* 65L09, 34A55.

The corresponding eigenvalues satisfy the asymptotic formula [**3**]

$$\lambda_n = \pi^2 n^2 + 2(h_1 - h_0) + \int_0^1 q(s)\,\mathrm{d}s + a_n\,, \tag{2.3}$$

where $(a_n) \in l^2$. It is a classical result [**1, 4**], that the potential q is uniquely determined by two sequences of eigenvalues corresponding to boundary conditions $(h_0 h_1)$ and $(h_0 h_2)$ with $h_1 \neq h_2$. Moreover it can be shown, that those sequences also uniquely determine the boundary conditions [**5**].

Therefore, two sequences of eigenvalues contain all information necessary to recover the potential as well as the corresponding boundary conditions. For notational convenience we write the parameters of the two Sturm-Liouville problems as vectors

$$\boldsymbol{q_1} := (h_0, h_1, q)\,, \quad \boldsymbol{q_2} := (h_0, h_2, q)\,,$$

$\lambda_{\boldsymbol{q}_i,n}$ for the n-th eigenvalue of problem $\boldsymbol{q}_i$, and

$$\boldsymbol{q} := (h_0, h_1, h_2, q)$$

for the full problem.

Now we define a least squares functional on the eigenvalues, which has the solution of the inverse problem as zero.

Definition 2.1. Suppose we are given (partial) spectral data $\lambda_{\boldsymbol{Q}_i,n}$ with (i,n) in $I \subseteq \{1,2\} \times \mathbb{N}$ of an unknown Sturm-Liouville problem $\boldsymbol{Q} = (H_0, H_1, H_2, Q)$. For a trial problem $\boldsymbol{q}$ and positive weights $\omega_{i,n}$, we define the functional

$$G(\boldsymbol{q}) := \sum_{(i,n)\in I} \omega_{i,n}(\lambda_{\boldsymbol{q}_i,n} - \lambda_{\boldsymbol{Q}_i,n})^2\,. \tag{2.4}$$

We note that $G(\boldsymbol{q})$ is positive, and zero if and only if both given sequences of eigenvalues match those of $\boldsymbol{q}$. If we have full knowledge of the two sequences $\lambda_{\boldsymbol{Q}_i,n}, (i,n) \in \{1,2\} \times \mathbb{N}$, this determines $\boldsymbol{q}$ uniquely, and hence $\boldsymbol{q} = \boldsymbol{Q}$.

To find such a $\boldsymbol{q}$, we minimize the functional with a conjugate gradient descent algorithm. First, for numerical stability it is good to know, that for each pair of interlacing sequences,

$$\lambda_{1,n} < \lambda_{2,n} < \lambda_{1,n+1} \qquad \text{or} \qquad \lambda_{2,n} < \lambda_{1,n} < \lambda_{2,n+1}\,,$$

which satisfy the asymptotics (2.3), there is a $\boldsymbol{q}$, with $G(\boldsymbol{q}) = 0$ [**5**].

The gradient of $\lambda_{\boldsymbol{q}_i,n}$ wrt. $\boldsymbol{q}_i$ is [**3**]

$$\nabla\lambda_{\boldsymbol{q}_i,n} = \begin{pmatrix} \frac{\partial\lambda_{\boldsymbol{q}_i,n}}{\partial h_0} \\ \frac{\partial\lambda_{\boldsymbol{q}_i,n}}{\partial h_1} \\ \frac{\partial\lambda_{\boldsymbol{q}_i,n}}{\partial h_2} \\ \frac{\partial\lambda_{\boldsymbol{q}_i,n}}{\partial q} \end{pmatrix} = \begin{pmatrix} -g^2_{\boldsymbol{q}_i,n}(0) \\ g^2_{\boldsymbol{q}_i,n}(1)\delta_{i,1} \\ g^2_{\boldsymbol{q}_i,n}(1)\delta_{i,2} \\ g^2_{\boldsymbol{q}_i,n}(x) \end{pmatrix}\,,$$

where $g_{\boldsymbol{q}_i,n}$ denotes the eigenfunction corresponding to $\lambda_{\boldsymbol{q}_i,n}$ with $\|g_{\boldsymbol{q}_i,n}\| = 1$.

It follows that the gradient of the functional is given by

$$\nabla G(\boldsymbol{q}) = 2\sum_{(i,n)\in I} \omega_{i,n}(\lambda_{\boldsymbol{q}_i,n} - \lambda_{\boldsymbol{Q}_i,n})\nabla\lambda_{\boldsymbol{q}_i,n}\,,$$

if $(\omega_{i,n})$ is summable.

Theorem 4.2 below shows that the gradients $\nabla\lambda_{\boldsymbol{q}_i,n}$ are linearly independent in $\mathbb{R}^3 \times L^2(0,1)$. This immediately implies the essential convexity of the functional:

THEOREM 2.2. *Let $\boldsymbol{q}_1 = (h_0, h_1, q)$ and $\boldsymbol{q}_2 = (h_0, h_2, q)$ be two Sturm-Liouville problems with $h_1 \neq h_2$. If I is finite or $(\omega_{i,n})$ is summable, the functional G has no local minima at $\boldsymbol{q}$ with $G(\boldsymbol{q}) > 0$, i.e.*

$$\nabla G(\boldsymbol{q}) = 0 \Longleftrightarrow G(\boldsymbol{q}) = 0\,.$$

Thus a conjugate gradient algorithm will not get trapped in local minima, as we will also observe in the examples.

3. Numerical Examples

We use the standard Polak-Ribiere conjugate gradient descent algorithm [**8**] to approximate the gradient flow and thus minimize the functional. To give the basic idea, we explain the simpler steepest descent:

(i) choose initial potential and boundary conditions $\boldsymbol{q}^{(0)} = (h_0^{(0)}, h_1^{(0)}, h_2^{(0)}, q^{(0)})$
(ii) while $G(\boldsymbol{q}^{(j)})$ too big do
  (a) compute the gradient $\nabla G(\boldsymbol{q}^{(j)})$
  (b) minimize the one dimensional function $G(\boldsymbol{q}^{(j)}) - \alpha \nabla G(\boldsymbol{q}^{(j)})$ wrt. α
  (c) set $\boldsymbol{q}^{(j+1)}$ equal to the minimizing potential

This straight forward minimization scheme has a major disadvantage: consecutive gradients are always orthogonal. To avoid this, conjugate gradient descent computes the direction for the one dimensional minimization using the current and previous gradients.

Note that the boundary points $g_{\boldsymbol{q}_i,n}(0)$ and $g_{\boldsymbol{q}_i,n}(1)$, needed in the computation of the gradient $\nabla \lambda_{\boldsymbol{q}_i,n}$, can be computed in a numerically well behaved way. Given the eigenvalue $\lambda_{\boldsymbol{q}_i,n}$, we can compute a multiple of the eigenfunction by solving an initial value problem. The value of the eigenfunction at the boundary then is just the (exactly known) initial value divided by the L^2 norm of the initial value solution.

We first apply the algorithm to a popular non-continuous potential [**2, 9**]

$$Q(x) = (7x - 0.7)\chi_{(0.1,0.3]}(x) + (3.5 - 7x)\chi_{(0.3,0.5]}(x) + 4\chi_{(0.7,0.9]}(x) + 2\chi_{(x,0.9,1]}(x)\,,$$

and choose

$$\boldsymbol{Q} = (3, 3, 0, Q) \qquad \text{and} \qquad \boldsymbol{q}^{(0)} = (2, 4, -1, 0) \tag{3.1}$$

with $I = \{1, 2\} \times \{0, \ldots, 29\}$ and weights $\omega_{i,n} = 1$. This will be our default setting, unless noted otherwise. In the figures we write $\Delta_2 = \|q - Q\|_2$ for the L^2 error of the reconstruction and $bc = (h_0; h_1; h_2)$ for the boundary conditions of the current approximation.

The results (figure 1) are comparable to the alternative boundary condition example with given boundary conditions in [**9**]. We get quite good results at 150 iterations, which keep getting better as we minimize the functional.

From around 150 iterations the boundary conditions stay almost constant. This suggests to reset the trial potential $q^{(j)}$ to zero at some iteration number j, while keeping the boundary conditions. In other words, we reset the forth component of $\boldsymbol{q}^{(j)} = (h_0^{(j)}, h_1^{(j)}, h_2^{(j)}, q^{(j)})$, and keep the others fixed. In figure 2 we indeed attain a significantly better approximation by setting the potential to zero a couple of times. In all our examples this was a very useful strategy to get faster convergence. But it is just heuristics – we do not really know how to choose the optimal number of iterations j. In practice, we wait until the boundary conditions stabilize and then set the potential to zero. This can be repeated until the convergence speed of the functional G does not improve any more.

The graph in figure 3 demonstrates, that the reconstruction of a smooth potential, using the same boundary conditions and number of eigenvalues, yields more accurate results.

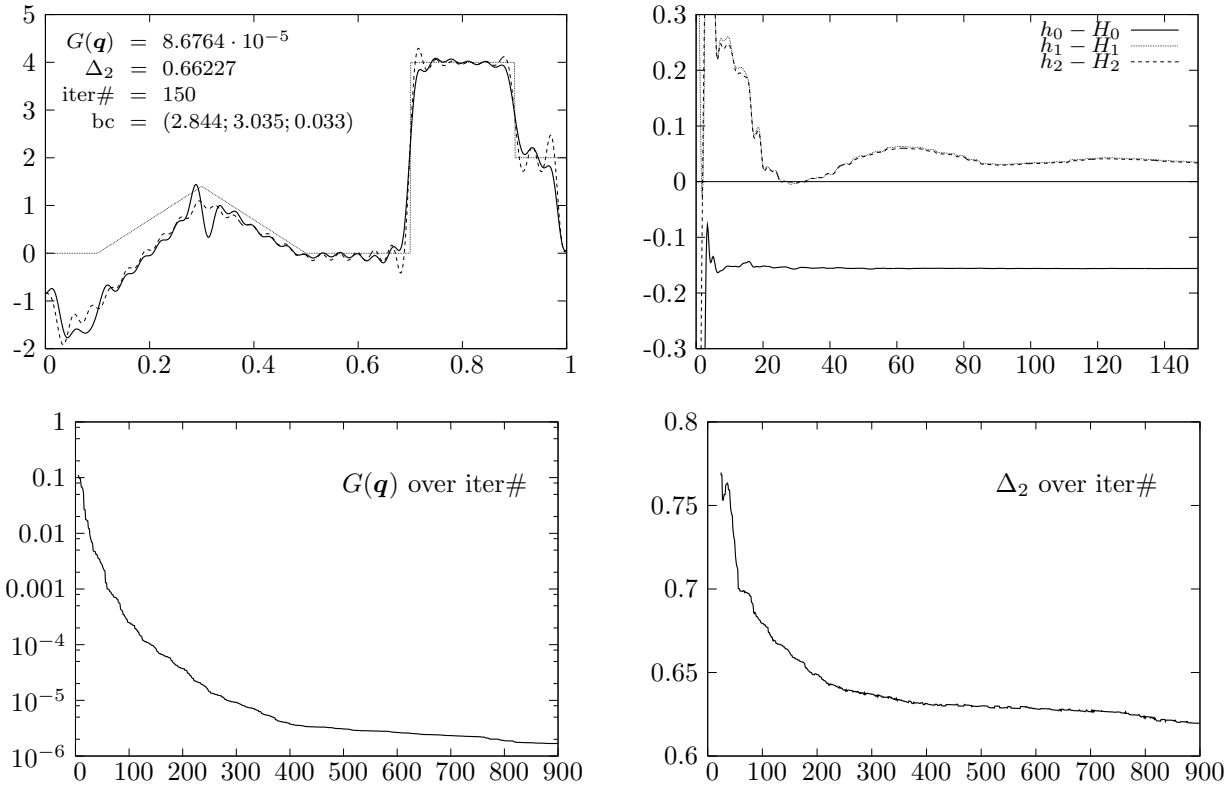

FIGURE 1. Graph of $\boldsymbol{q}^{(150)}$, $\boldsymbol{q}^{(840)}$ (light) and boundary conditions, $G(\boldsymbol{q})$, and L^2 error versus the number of iterations.

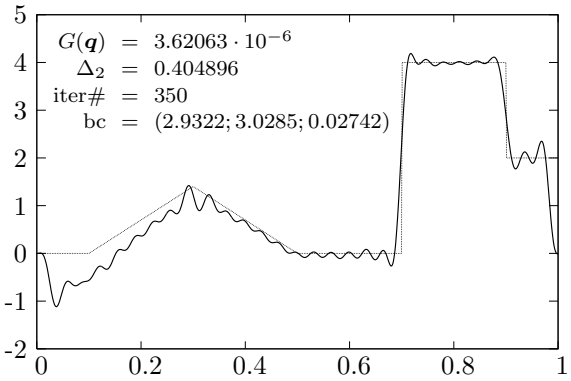

FIGURE 2. Above example with setting $q^{(30)} = q^{(60)} = q^{(110)} = 0$.

This overall behavior is also true for worse guesses of the initial boundary conditions. If we take for example

$$\boldsymbol{q}^{(0)} = (0, 0, 0, 0)\,,$$

the boundary conditions converge slowly, but steadily (figure 4). Again, setting $q^{(j)} = 0$ a couple of times increases the speed of convergence dramatically.

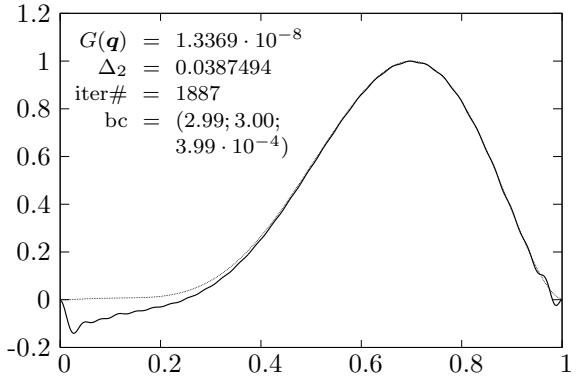

FIGURE 3. Approximation of a smooth potential.

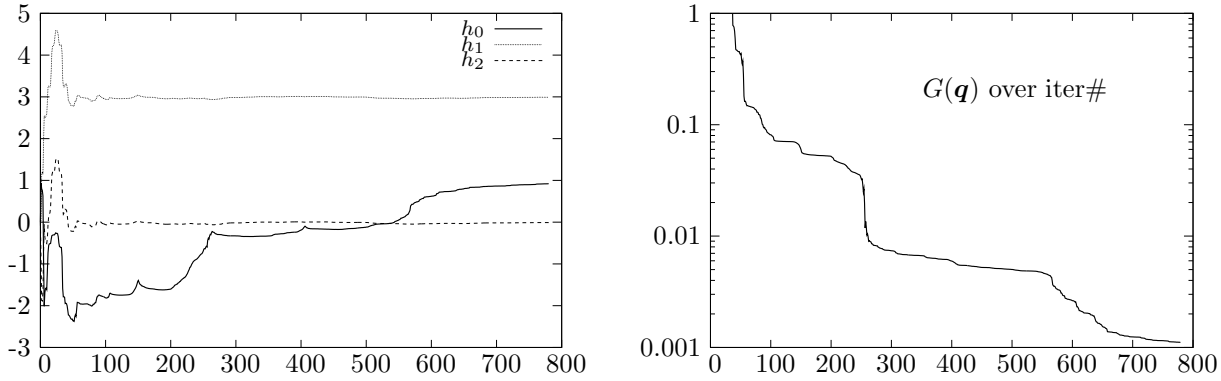

FIGURE 4. The convergence of the boundary conditions and of the functional for initial problem $\boldsymbol{q}^{(0)} = (0, 0, 0, 0)$.

In the case of noisy data, the algorithm is of course much more unstable than the version with fixed boundary conditions [**9**]. For the following examples, we add random noise $|\tilde{\lambda}_{\boldsymbol{Q}_i,n} - \lambda_{\boldsymbol{Q}_i,n}| \leq r$ to the eigenvalues. To see the limitations of this approach, we first use $r = 0.1$ and $H_0 = h_0 = H_1 = h_1 = 3$, $H_2 = h_2 = 0$ (figure 5), i.e. we already start with the correct boundary conditions.

The algorithm passes through a potential, which is reasonably close to the original potential Q. But from there, the steepest descent leads to a potential, which is not in any way similar to the one we want to recover. In the graph of the L^2 error, we see that there are roughly 20 iterations of good approximations. Afterwards the approximations quickly get worse than our initial guess.

Yet, for smaller errors in the eigenvalues, this effect is less dramatic. Setting $r = 0.01$ and using the problem (3.1), we get good approximations for around 100 iterations (figure 6). The L^2 error also rises more slowly than for the case with larger noise level.

Setting $q^{(89)} = 0$ again improves the performance significantly (figure 7). We get reasonable approximations for all 400 iterations but the first 20 after each setting the potential to zero.

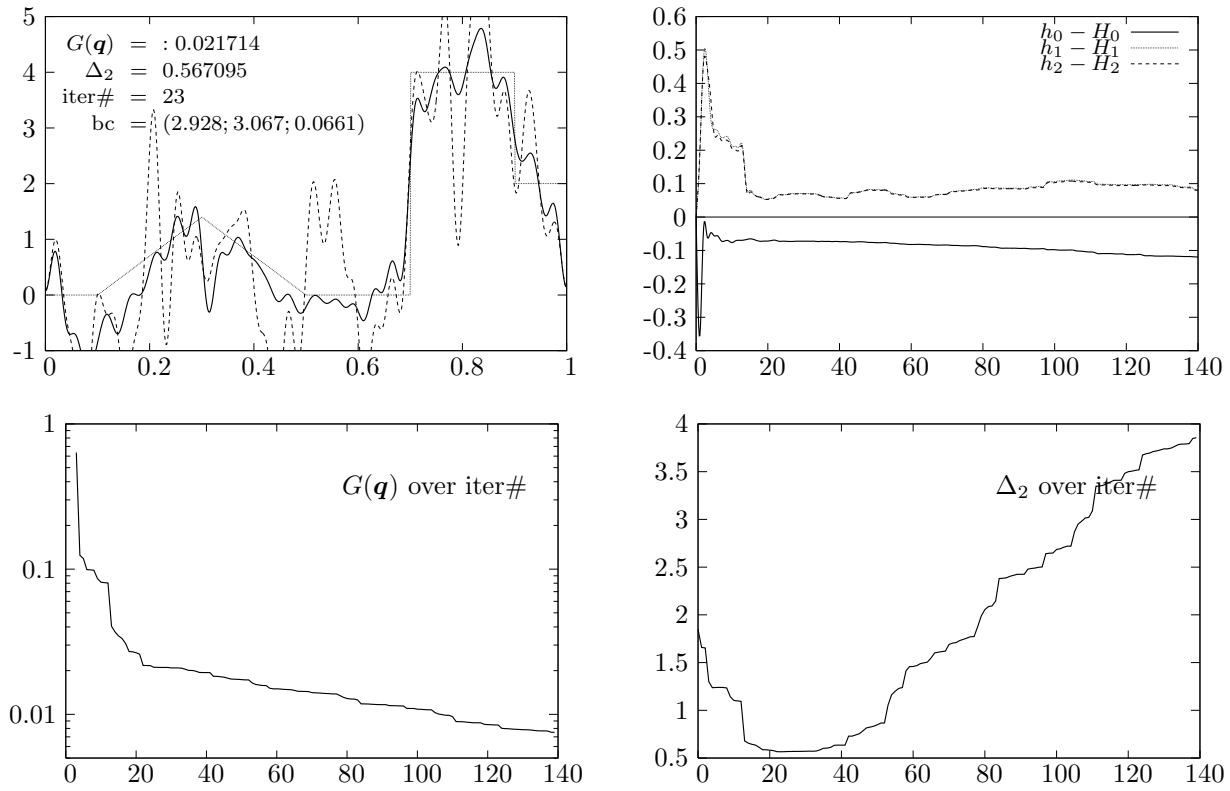

FIGURE 5. Example with errors of magnitude $r = 0.1$. The best L^2 approximation at 23 iterations and $q^{(55)}$.

4. Linear Independence of the Gradients

First, we borrow the central lemma of the independence proof of the original functional [**9**]. To that end, we define the Wronskian $[f, g] = fg' - f'g$ and the bilinear form

$$\begin{array}{rccc} \Gamma: & H^1([0,1],\mathbb{R})^2 & \longrightarrow & \mathbb{R} \\ & (f,g) & \mapsto & \int_0^1 [f,g]\,\mathrm{d}x \end{array},$$

which is bounded by

$$|\Gamma(f,g)| \leq \|f\|_{H_1}\|g\|_{H_1}, \text{ i.e. } \|\Gamma(f,\cdot)\| = \|f\|_{H_1}.$$

(We use the definition $\|f\|_{H^1} = \sqrt{\|f\|_{L^2}^2 + \|f'\|_{L^2}^2}$ with distributional derivatives.)

Let $s_{i,n,q}$ and $c_{i,n,q}$ be the solutions of the differential equation (SL) for the eigenvalue parameter $\lambda_{\boldsymbol{q}_i,n}$ and initial values

$$\begin{array}{rclcrcl} s_{i,n,q}(1) & = & 1\,, & & c_{i,n,q}(1) & = & 1\,, \\ s'_{i,n,q}(1) & = & -h_1\,, & & c'_{i,n,q}(1) & = & -h_2\,. \end{array}$$

LEMMA 4.1 ([**9**]). *Given two Sturm-Liouville problems $\boldsymbol{q}_1 = (h_0, h_1, q)$ and $\boldsymbol{q}_2 = (h_0, h_2, q)$ with $h_1 \neq h_2$, we have*

$$\Gamma(c_{i,n,q}s_{i,n,q}, g^2_{\boldsymbol{q}_j,m}) = (-1)^i(h_2 - h_1)\delta_{n,m}\delta_{i,j} \qquad \textit{for all } i,j \in \{1,2\} \textit{ and } m,n \in \mathbb{N}$$

for the normalized eigenfunctions $g_{\boldsymbol{q}_i,n}$ and $s_{i,n,q}$, $c_{i,n,q}$ as defined above.

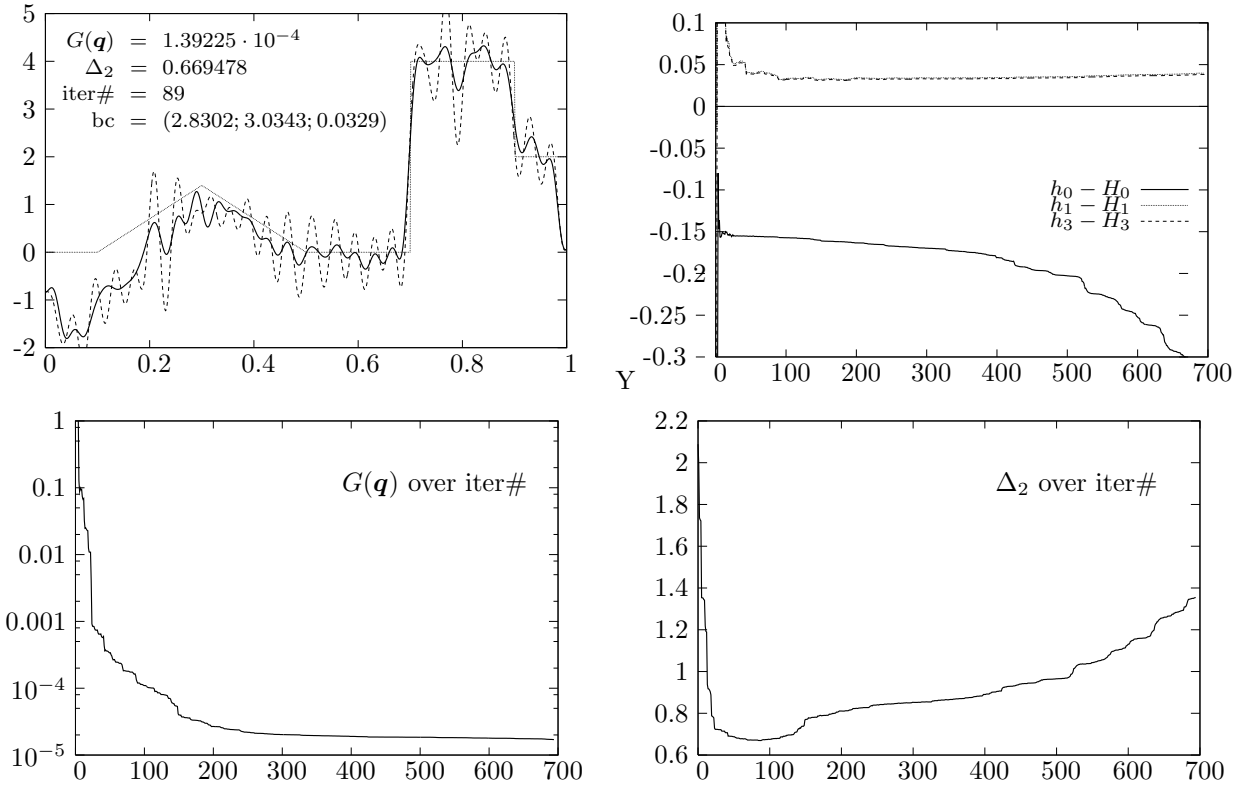

FIGURE 6. Example with errors of magnitude $r = 0.01$. The best L^2 approximation at 89 iterations and $q^{(300)}$.

Using the alternative bilinear form

$$\begin{array}{rccl} \tilde{\Gamma}: & H^1([0,1],\mathbb{R}) \times (\mathbb{R}^3 \times L^2([0,1],\mathbb{R})) & \longrightarrow & \mathbb{R} \\ & (f,(a,b,c,g)) & \mapsto & -2\int_0^1 f'g\,\mathrm{d}x + f(1)b + f(1)c + f(0)a \end{array},$$

and integration by parts

$$\Gamma(f,g) = -2\int_0^1 f'g\,\mathrm{d}x + fg(1) - fg(0)\,,$$

we get the corresponding statement for the gradients $\lambda_{\boldsymbol{q}_j,m}$

$$\tilde{\Gamma}(c_{i,n,q}s_{i,n,q}, \nabla\lambda_{\boldsymbol{q}_j,m}) = \Gamma(c_{i,n,q}s_{i,n,q}, g^2_{\boldsymbol{q}_j,m}) = (-1)^i(h_2 - h_1)\delta_{n,m}\delta_{i,j}\,.$$

Finally, since $\tilde{\Gamma}$ is bounded by

$$|\tilde{\Gamma}(f,(a,b,c,g))| \le 2\|f\|_{H^1}\|g\|_{L^2} + \sqrt{2}\|f\|_{H^1}(|a|+|b|+|c|) \le 2\|f\|_{H^1}(\|g\|_{L^2} + |a| + |b| + |c|)$$

and, in particular, continuous in the 2nd component, we immediately get the linear independence theorem.

THEOREM 4.2. *With the notations of the above lemma, the set of gradients of the eigenvalues*

$$\left\{\nabla\lambda_{\boldsymbol{q}_i,n} | (i,n) \in \{1,2\} \times \mathbb{N}\right\}$$

is linearly independent in $\mathbb{R}^3 \times L^2$.

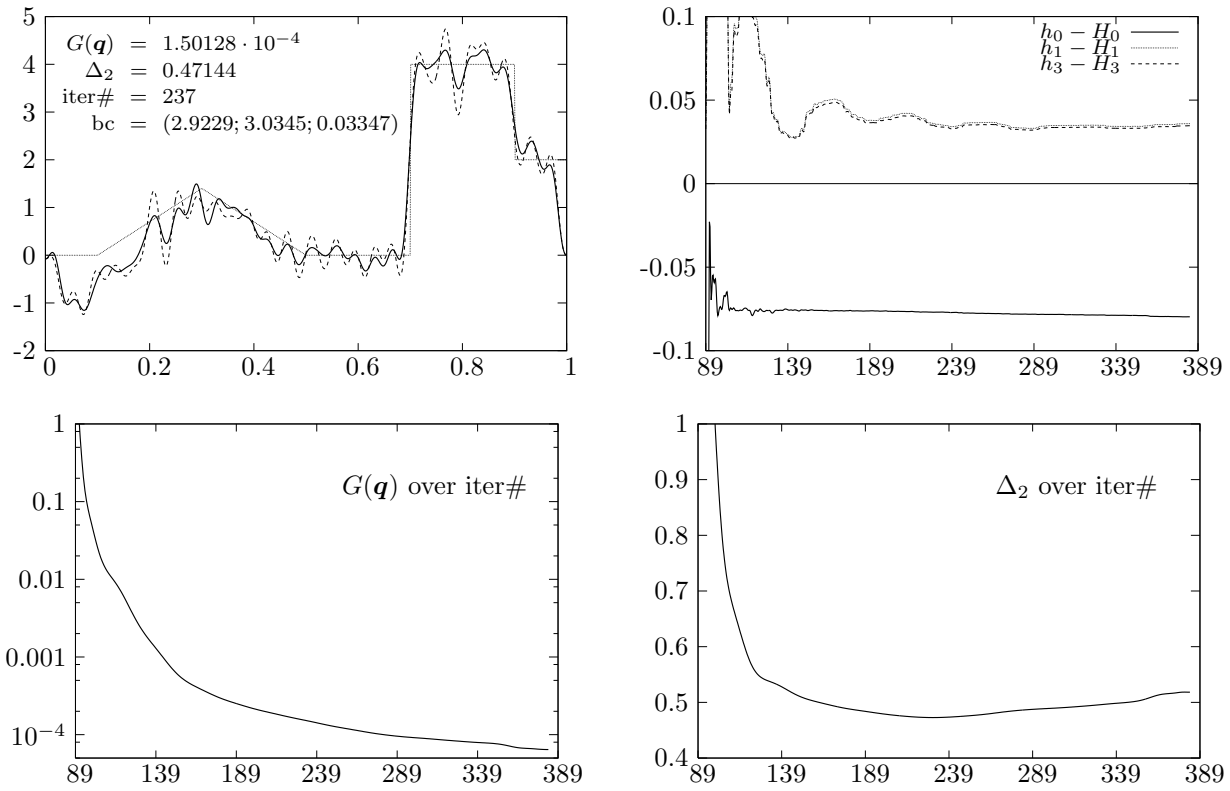

FIGURE 7. Example with errors of magnitude $r = 0.01$ and $\boldsymbol{q}^{(89)}$ set to zero. The best L^2 approximation at 237 iterations and $q^{(389)}$.

PROOF. Suppose for some fixed (i, n) we have

$$\nabla\lambda_{\boldsymbol{q}_i,n} = \sum_{k\in\mathbb{N}} a_k \nabla\lambda_k$$

in $\mathbb{R}^3 \times L^2$, where $a_k \in \mathbb{R}$ and $\nabla\lambda_k = \nabla\lambda_{\boldsymbol{q}_{j_k},m_k}$ with $(j_k, m_k) \neq (i, n)$. But this would imply

$$(-1)^i(h_2 - h_1) = \tilde{\Gamma}(c_{i,n,q}s_{i,n,q}, \nabla\lambda_{\boldsymbol{q}_i,n}) = \tilde{\Gamma}\left(c_{i,n,q}s_{i,n,q}, \sum_{k\in\mathbb{N}} a_k\nabla\lambda_k\right) = \sum_{k\in\mathbb{N}} \tilde{\Gamma}(c_{i,n,q}s_{i,n,q}, a_k\nabla\lambda_k) = 0$$

□

References

1. G. Borg, *Eine Umkehrung der Sturm-Liouvilleschen Eigenwertaufgabe. Bestimmung der Differentialgleichung durch die Eigenwerte*, Acta Math. **78** (1946), 1–96.
2. B. M. Brown, V. S. Samko, I. W. Knowles, and M. Marletta, *Inverse spectral problem for the Sturm-Liouville equation*, Inverse Problems **19** (2003), no. 1, 235–252.
3. E. L. Isaacson and E. Trubowitz, *The inverse Sturm-Liouville problem. I*, Comm. Pure Appl. Math. **36** (1983), no. 6, 767–783.
4. N. Levinson, *The inverse Sturm-Liouville problem*, Mat. Tidsskr. B. **1949** (1949), 25–30.
5. B. M. Levitan, *Inverse Sturm-Liouville problems*, VSP, Zeist, 1987, Translated from the Russian by O. Efimov.
6. B. D. Lowe, M. Pilant, and W. Rundell, *The recovery of potentials from finite spectral data*, SIAM J. Math. Anal. **23** (1992), no. 2, 482–504.

7. J. R. McLaughlin, *Analytical methods for recovering coefficients in differential equations from spectral data*, SIAM Rev. **28** (1986), no. 1, 53–72.
8. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C*, second ed., Cambridge University Press, Cambridge, 1992.
9. N. Roehrl, *A least squares functional for solving inverse Sturm-Liouville problems*, Inverse Problems **21** (2005), 2009–2017.
10. W. Rundell and P. E. Sacks, *Reconstruction techniques for classical inverse Sturm-Liouville problems*, Math. Comp. **58** (1992), no. 197, 161–183.

Fachbereich Mathematik, Univerität Stuttgart, 70550 Stuttgart, Germany
E-mail address: norbert.roehrl@mathematik.uni-stuttgart.de